

*Please provide the following information, and submit to the NOAA DM Plan Repository.*

**Reference to Master DM Plan (if applicable)**

*As stated in Section IV, Requirement 1.3, DM Plans may be hierarchical. If this DM Plan inherits provisions from a higher-level DM Plan already submitted to the Repository, then this more-specific Plan only needs to provide information that differs from what was provided in the Master DM Plan.*

URL of higher-level DM Plan (if any) as submitted to DM Plan Repository:

**1. General Description of Data to be Managed****1.1. Name of the Data, data collection Project, or data-producing Program:**

Assessing cryptic reef diversity of colonizing marine invertebrates using Autonomous Reef Monitoring Structures (ARMS) deployed at coral reef sites in Timor-Leste from 2012 to 2014

**1.2. Summary description of the data:**

The data described here, including photographs, genetic sequences, and specimen information, were collected by the NOAA Coral Reef Ecosystem Program (CREP) from Autonomous Reef Monitoring Structures, or ARMS, moored for two years at fixed climate survey sites located on hard bottom shallow water (< 15 m) habitats in Timor-Leste. Climate sites were established in Timor-Leste in October 2012 to establish ecological baselines for climate change by measuring multiple features of the coral reef environment (in addition to the data described herein) over time.

Three ARMS units were typically deployed by SCUBA divers at each survey site. Each ARMS unit, constructed in-house by CREP, consisted of 23 cm x 23 cm gray, type 1 PVC plates stacked in alternating series of 4 open and 4 obstructed layers and attached to a base plate of 35 cm x 45 cm, which was affixed to the reef. Upon recovery, each ARMS unit was encapsulated, brought to the surface, and disassembled and processed. Disassembled plates were photographed to document recruited sessile organisms, scraped clean and preserved in 95% ethanol for DNA processing. Recruited motile organisms were sieved into 3 size fractions: 2 mm, 500 µm, and 100 µm. The 500 µm and 100 µm fractions were bulked and also preserved in 95% ethanol for DNA processing. The 2 mm fraction was sorted into morphospecies, photographed, and identified to the lowest taxonomic identification possible. The plate photographs, sequences generated from the DNA metabarcoding of the scrapings and the 500- and 100-µm fractions, specimen photographs, and specimen identifications are included in the ARMS dataset. The data can be accessed online via the NOAA National Centers for Environmental Information (NCEI) Ocean Archive.

ARMS are used by CREP to assess and monitor cryptic reef diversity across the Pacific. Developed in collaboration with the Census of Marine Life (CoML) Census of Coral Reef Ecosystems (CReefs), ARMS are designed to mimic the structural complexity of a reef

and attract/collect colonizing marine invertebrates. The key innovation of the ARMS method is biodiversity is sampled over precisely the same surface area in the exact same manner. Thus, the use of ARMS is a systematic, consistent, and comparable method for monitoring the marine cryptobiota community over time.

**1.3. Is this a one-time data collection, or an ongoing series of measurements?**

One-time data collection

**1.4. Actual or planned temporal coverage of the data:**

2012-10-15 to 2014-10-09

**1.5. Actual or planned geographic coverage of the data:**

W: 125.01319, E: 127.31222, N: -8.22438, S: -8.85321

Extent of climate survey sites in Timor-Leste.

**1.6. Type(s) of data:**

*(e.g., digital numeric data, imagery, photographs, video, audio, database, tabular data, etc.)*

Table (digital), image (digital)

**1.7. Data collection method(s):**

*(e.g., satellite, airplane, unmanned aerial system, radar, weather station, moored buoy, research vessel, autonomous underwater vehicle, animal tagging, manual surveys, enforcement activities, numerical model, etc.)*

Instrument: Not applicable

Platform: Not applicable

Physical Collection / Fishing Gear: Autonomous Reef Monitoring Structures (ARMS)

**1.8. If data are from a NOAA Observing System of Record, indicate name of system:**

**1.8.1. If data are from another observing system, please specify:**

**2. Point of Contact for this Data Management Plan (author or maintainer)**

**2.1. Name:**

Annette M DesRochers

**2.2. Title:**

Metadata Contact

**2.3. Affiliation or facility:**

**2.4. E-mail address:**

annette.desrochers@noaa.gov

**2.5. Phone number:**

(808)725-5461

### 3. Responsible Party for Data Management

*Program Managers, or their designee, shall be responsible for assuring the proper management of the data produced by their Program. Please indicate the responsible party below.*

**3.1. Name:**

Molly A Timmers

**3.2. Title:**

Data Steward

### 4. Resources

*Programs must identify resources within their own budget for managing the data they produce.*

**4.1. Have resources for management of these data been identified?**

Yes

**4.2. Approximate percentage of the budget for these data devoted to data management (specify percentage or "unknown"):**

Unknown

### 5. Data Lineage and Quality

*NOAA has issued Information Quality Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information which it disseminates.*

**5.1. Processing workflow of the data from collection or acquisition to making it publicly accessible**

*(describe or provide URL of description):*

Lineage Statement:

Autonomous Reef Monitoring Structures (ARMS) field, laboratory, and analysis methods used for ARMS in Timor-Leste

Process Steps:

- FIELD METHOD: ARMS, composed of nine PVC plates (23 cm x 23 cm) stacked in alternating series of open and semi-enclosed layers, were affixed to the seafloor between 12–15 m in replicate sets of three. They remained on the benthos for two years during which time they were naturally colonized with marine organisms. After the 2-year deployment period, the ARMS units were encapsulated within a 106-µm nitex-lined crate, brought to the surface, placed within a large seawater holding bin and transported to shore. On shore, they were disassembled plate by plate, with both sides photo-documented. The plates were then scraped clear of all the accumulated sessile biomass and immediately homogenized in a blender, filtered with a 40-µm net, subsampled, and preserved for metabarcoding. The seawater used during processing was sieved using 2-mm, 500-µm and 106-µm geologic sieves to create three size fractions. The >2 mm fraction was sorted to morphospecies, photographed, and brachyuran crabs were preserved for DNA barcoding. The two smaller motile fractions were preserved for additional lab and

molecular processing. (Citation: PIFSC. 2017. Interdisciplinary baseline ecosystem assessment surveys to inform ecosystem-based management planning in Timor-Leste: Final Report. NOAA Pacific Islands Fisheries Science Center, PIFSC Special Publication, SP-17-02, 234p.)

- LAB METHOD: Decantation—Due to sediment within the 500- $\mu$ m and 106- $\mu$ m fractions that can inhibit metabarcoding laboratory processing, a decantation procedure was conducted on these fractions from each ARMS unit to separate the sediment from the organic matter. Upon the completion of the decantation process, half of the sample was crushed with a mortar and pestle for DNA extraction and metabarcoding while the other half was preserved as a backup. DNA barcoding—Legs from brachyuran crabs were subsampled, and genomic DNA was extracted using standard proteinase-k digestion followed by phenol-chloroform extraction on the AutoGenprep 965 (Autogen). Primers were used to target approximately 658 base pairs of the COI gene and automated sequencing techniques were used to sequence in both directions. DNA metabarcoding—DNA was extracted from 10 grams of the homogenized sessile scrapings and from the decanted 500- $\mu$ m and 100- $\mu$ m motile fractions using the MO-Bio PowerMax Soil extraction kits. Using the reverse primer, jgHCO2198, and the forward primer, mlCOIintF, a 313 base pair fragment of COI was amplified using a PCR (polymerase chain reaction) touchdown protocol with 16 initial cycles: denaturation for 10 seconds at 95°C, annealing for 30 seconds at 62°C (–1°C per cycle), and extension for 60 seconds at 72°C, followed by 25 cycles at 46°C annealing temperature. PCRs were performed in triplicates and inspected on agarose gels. Triplicate PCR products were pooled, cleaned using Agencourt AMPure beads, and quantified using Biotum AccuClear Ultra High Sensitivity Quantification Kit. PCR products were then inserted directly into the Kappa Systems Hyper-Prep sample kit using dual-end Illumina adapters for ligation. Sample libraries were validated by visualization on an Agilent 2100 BioAnalyzer, quantified using qPCR, pooled, and sequenced on an Illumina MiSeq platform. Each library yielded approximately 250,000 reads per sample, and a standard quality control filter was run to parse the Illumina reads into FASTQ files sorted by index. (Citation: PIFSC. 2017. Interdisciplinary baseline ecosystem assessment surveys to inform ecosystem-based management planning in Timor-Leste: Final Report. NOAA Pacific Islands Fisheries Science Center, PIFSC Special Publication, SP-17-02, 234p.)

- ANALYSIS: Morphospecies (2-mm size fraction)—Overall abundance of >2 mm organisms was averaged between ARMS units recovered at each site to give a site-level metric. Organisms were additionally averaged by ARMS unit at the island scale for comparison with other ARMS recovery locations across the Pacific. Dominant phyla and taxa groups within phyla were averaged between ARMS units and compared across sites. Crab DNA barcoding—Resulting sequences of crabs were clustered into Operational Taxonomic Units (OTUs) and blasted (cross checked) against existing DNA-barcoding libraries (Barcode of Life Data Systems [BOLD] and Moorea Biocode). Matched sequences with >97% identity and >85% coverage were identified to an existing record of the species within the databases. Those crab sequences with <97% identity and >85% coverage underwent a phylogenetic

Bayesian approach using the Statistical Assignment Package (SAP) to assign OTUs to higher taxonomic levels in the absence of a direct match. Species richness was averaged by ARMS unit at each site and examined on the island scale in relation to the richness of brachyuran crabs from other ARMS units collected by NOAA-CREP in the Pacific Ocean. Broad scale richness values were calculated per ARMS unit richness rather than by island due to the variability in the number of ARMS units deployed across islands. Metabarcoding bioinformatics—Sequences were assembled, trimmed, cleaned, and dereplicated following standard bioinformatics techniques using available software programs. Dereplicated sequences were then aligned to COI barcodes from the BOLD database. Matched sequences =97% identity and =85% coverage are presented herein. Sequences that did not have a direct match have not been directly DNA barcoded and thus species resolution is not available. Once the phylogenetic approaches and bioinformatic software have been refined, the remaining unknown sequences can be determined. Currently available software is not capable of working through 10 million plus sequence reads that span across multiple phyla. However, through the efforts of a third-party bioinformation specialist working on these data sets for Timor-Leste, a solution will be found in the near future to provide phyla-based resolution of the remaining sequences that will indicate percent cover of the phyla communities that have recruited to the ARMS units. (Citation: PIFSC. 2017. Interdisciplinary baseline ecosystem assessment surveys to inform ecosystem-based management planning in Timor-Leste: Final Report. NOAA Pacific Islands Fisheries Science Center, PIFSC Special Publication, SP-17-02, 234p.)

**5.1.1. If data at different stages of the workflow, or products derived from these data, are subject to a separate data management plan, provide reference to other plan:**

**5.2. Quality control procedures employed (describe or provide URL of description):**

The data entered in the MS Access database is quality controlled following data entry.

## **6. Data Documentation**

*The EDMC Data Documentation Procedural Directive requires that NOAA data be well documented, specifies the use of ISO 19115 and related standards for documentation of new data, and provides links to resources and tools for metadata creation and validation.*

**6.1. Does metadata comply with EDMC Data Documentation directive?**

Yes

**6.1.1. If metadata are non-existent or non-compliant, please explain:**

**6.2. Name of organization or facility providing metadata hosting:**

NMFS Office of Science and Technology

**6.2.1. If service is needed for metadata hosting, please indicate:****6.3. URL of metadata folder or data catalog, if known:**

<https://www.fisheries.noaa.gov/inport/item/46159>

**6.4. Process for producing and maintaining metadata**

*(describe or provide URL of description):*

Metadata produced and maintained in accordance with the NOAA Data Documentation Procedural Directive: [https://nosc.noaa.gov/EDMC/DAARWG/docs/EDMC\\_PD-Data\\_Documentation\\_v1.pdf](https://nosc.noaa.gov/EDMC/DAARWG/docs/EDMC_PD-Data_Documentation_v1.pdf)

**7. Data Access**

*NAO 212-15 states that access to environmental data may only be restricted when distribution is explicitly limited by law, regulation, policy (such as those applicable to personally identifiable information or protected critical infrastructure information or proprietary trade information) or by security requirements. The EDMC Data Access Procedural Directive contains specific guidance, recommends the use of open-standard, interoperable, non-proprietary web services, provides information about resources and tools to enable data access, and includes a Waiver to be submitted to justify any approach other than full, unrestricted public access.*

**7.1. Do these data comply with the Data Access directive?**

Yes

**7.1.1. If the data are not to be made available to the public at all, or with limitations, has a Waiver (Appendix A of Data Access directive) been filed?****7.1.2. If there are limitations to public data access, describe how data are protected from unauthorized access or disclosure:****7.2. Name of organization of facility providing data access:**

National Centers for Environmental Information - Silver Spring, Maryland (NCEI-MD)

**7.2.1. If data hosting service is needed, please indicate:****7.2.2. URL of data access service, if known:**

<http://accession.nodc.noaa.gov/0169338>

<http://accession.nodc.noaa.gov/0169338>

<http://accession.nodc.noaa.gov/0169338>

<http://accession.nodc.noaa.gov/0169338>

**7.3. Data access methods or services offered:**

Data can be accessed online via the NOAA National Centers for Environmental Information (NCEI) Ocean Archive.

**7.4. Approximate delay between data collection and dissemination:**

Unknown

**7.4.1. If delay is longer than latency of automated processing, indicate under what authority data access is delayed:**

## **8. Data Preservation and Protection**

*The NOAA Procedure for Scientific Records Appraisal and Archive Approval describes how to identify, appraise and decide what scientific records are to be preserved in a NOAA archive.*

### **8.1. Actual or planned long-term data archive location:**

*(Specify NCEI-MD, NCEI-CO, NCEI-NC, NCEI-MS, World Data Center (WDC) facility, Other, To Be Determined, Unable to Archive, or No Archiving Intended)*

NCEI\_MD

#### **8.1.1. If World Data Center or Other, specify:**

#### **8.1.2. If To Be Determined, Unable to Archive or No Archiving Intended, explain:**

### **8.2. Data storage facility prior to being sent to an archive facility (if any):**

Pacific Islands Fisheries Science Center - Honolulu, HI

### **8.3. Approximate delay between data collection and submission to an archive facility:**

Unknown

### **8.4. How will the data be protected from accidental or malicious modification or deletion prior to receipt by the archive?**

*Discuss data back-up, disaster recovery/contingency planning, and off-site data storage relevant to the data collection*

NOAA IRC and NOAA Fisheries ITS resources and assets. The MS Access database is stored on the PIFSC network and regularly backed up by ITS.

## **9. Additional Line Office or Staff Office Questions**

*Line and Staff Offices may extend this template by inserting additional questions in this section.*